

Fig.1

C2GnT3

1 M K I F K C Y F K H T L O O K V F I L F L T L W L L S L L K L L N V R R L F P Q 40
 1 ATGAGATATTCAAAATGTTATTTAAACATACCCCTACAGCAGAAAGTTTTCATCCTGTTTAAACCTTCTTAAGCTTCTTAATGTGAGACGACTCTTTCCGCRA 120
 41 K D I Y L V E Y S L S T S P F V R N R Y T H V K D E V R Y E V N C S G I Y E Q E 80
 121 AAAGACATTTACTGGTTGAGTACTCCCTAAGTACCTCGCTTTTGTGAAGAACAGATACACTCATGTTAAGGATGAAGTCAGGTGAAGTAACTGTTCCGGTATCTATGAACAGGAG 240
 81 P L E I G K S L E I R R R D I I D L E D D V V A M T S D C D I Y Q T L R G Y A 120
 241 CCTTTGGAAATTCGAAAGAGTCTGGAAATAAGAAGAGGACATCATTTGAGGATGATGTTGTGGCAATGACCACTGATTTGACATTTATCAGACTCTAAGAGGTTATGCT 360
 121 Q K L V S K E E K S F P I A Y S L V V H K D A I M V E R L I H A I Y N Q H N I Y 160
 361 CHAAAGCTTCTCTCAAGGAGGAGAAAGCTTCCCAATAGCCTATTTCTTTGGTTGTCACAAAGATGCAATATGTTGAAGGCTTATCCATGCTATATACACAGCAGCAATATTTAC 480
 161 C I H Y D R K A P D T F K V A M N N L A K C F S N I F I A S K L E A V E Y A H I 200
 481 TGCATCCATTTATGATCGTAAGGCACCTGATACCTTCAAGTCTTCAAGTCCATCCAGTGAATATGTTATCAACTGTGTGGCAAGATTTCCCTGAACTCAATTTTGAATTTGGTG 240
 201 S R L Q A D L N C L S D L L K S S I Q W K Y V I N L C G Q D F P L K S N F E L V 720
 601 TCCAGACTCCAGGCTGATTAAATTTGCTTGTGGACCTTCTGAGTCTTCAAGTCCATCCAGTGAATATGTTATCAACTGTGTGGCAAGATTTCCCTGAACTCAATTTTGAATTTGGTG 280
 241 S E L K K L N G A N M L E T V K P P N S K L E R F T Y H H E L R R V P Y E Y V K 840
 721 TCAGAGTTGAAAAAACAATCAATGGAGCAATATGTTGGACGCTGAACCCCAACAGTAAATGSAAGATTCACCTACCATCATGAAGTGTAGACGGTGCCTTATGAATATGTGAAG 320
 281 L P I R T N I S K E A P P H N I Q I F V G S A Y F V L S Q A F V K Y I F N S I 960
 841 CTACCAATAAGGACAAACATCTCCAAGGAGCACCCTCCATACATTTCTGACATTTTCTGACGTCTTATTTTGAAGTCAAGCATTTGTAAATATATTTTCAACACTCCATC 360
 321 V Q D F F A W S K D T Y S P D E H F W A T L I R V P G I P G E I S R S A Q D V S 1080
 961 GTTCAAGACTTTTTTGCTGCTCTAAGACACATATCTCTCATGAGCACTTTTGGCTACCTTGATTCGGTTCAGGAATACCTGGGAGATTTCCAGATCAGCCAGGATGTGTCT 400
 361 D L Q S K T R L V K W N Y Y E G F F Y P S C T G S H L R S V C I Y G A A E L R W 1200
 1081 GATCTGAGAGTAACTCGCTTGTCAAGTGGAAATTAATGAAAGCTTTTCTATCCAGTTGACTGGATCTCACCTTCGAAGCGTGTGTATTTATGGAGCTGCAGAAATTAAGGTGG 440
 401 L I K D G H W F A N K F D S K V D P I L I K C L A E K L E E Q Q R D W I T L P S 1320
 1201 CTTATCAAGATGGACATTTGCTTCTTAATTAATTTGATTCTAAGGTGGACCTATCTTGATTAATGCTTGGCAGAAAAGCTTGAAGAACAGCAGAGACTGGATCACTTTGCCCTCA 453
 441 E K L F M D R N L T T S * 1362
 1321 GAAAGTTATTTATGGATAGAAATCTCACTACCATCATGA

C2GnT3 : MKIFKCYFKHTLQKVFLEF-TLMVSLKLLNVRFPQNDIYLYEYSLSTSPFVRN-RYTHVKDEVRYEVCSCGIVEQEPL : 83
 C2GnT2 : -----MVQWKRIGLHYLWAGCMATVAIKLSFELKCDSDHGLGLESSESQOYCRNIDYFLKIPAKRSINCSVTFGDOEA : 80
 C2GnT1 : -----MLRTLRRLTSYPTKYFMVLISLITSVLRHOKPEYSVR-----H--TELAGENPSS-DINCTKALOCBYNE : 69
 IGnT : -----MPISMRYPFISVSVIIFIFSVFNEG-----GDSRQRLN-----ISDPLRLTQVCTSFING----- : 54

O*

C2GnT3 : IGKS-----LEIRRRDIIDLEDDVVAWTSDCDIYQILRGVAKLVSKSEKSFPTAYSLSLVVHKDAIMVERLIHAMVNOHNDVCIHY : 164
 C2GnT2 : VLOAIFNNLEWKKRPF-FITHTLSLTRDCEHKAERKEIQEFISKEEVEFPPTAYSMVHEKTIENEERLIRAVYAPONTLYCVHV : 164
 C2GnT1 : TOKVKLEILTAKFKRWIPDDINNTSDCSSEIKRRKYIVDEPISKEEAEFPPTAYSIVVHKEIEMLDRLIRAHYMPONFECVHV : 154
 IGnT : ---K-----TFLWKNKLMHEKSSCKEYLTQSHYIPAPISKEEAEFPPTAYSIVVHKEIEMLDRLIRAHYMPONFECVHV : 126

*

C2GnT3 : DRKADIFKVMNNLAKCFSENIETASKLEAVEYAHISRLQADLNC DLTKSSI QWKYVINC GODEPLKSNTELIVSELKKNKA : 249
 C2GnT2 : DEKSEEFKEAKAISCENEFASKIVRVVYASMSRQADLNCMEDLQSSPWKYFLNCTGTFPIKSNAMVQALKNLNGR : 249
 C2GnT1 : DTKSESYLAAMGASCFSNVVASRLESVYASMSRQADLNC KDLYAMSANWKYLINC GODEPLKSNTELIVSELKKNKA : 239
 IGnT : DEKATTEERDAVEQLISCENAFASKMEPVWVGGSRLQADLNC IRDLSAFEVSWKYVINC GODEPLKSNTELIVSELKKNKA : 211

*

C2GnT3 : NMIEIVKPPNSKLERFTYHHELRRVPYEVYKLPITNISKRAPEENIQIEVGSAYEVLISOAEVVKYIFNNSIVQDFFAMSKDTTSP : 334
 C2GnT2 : NSMSEVPPKHKEKWKYFEWARD-----TLHL-TNKKKDEPEENIMETGNAYIVASRDEVQHLKNPSOOLIEVVKDTTSP : 328
 C2GnT1 : NNEERMRPSHKEERWKRYENNG-----ALTN-TGIVKMLPELETPLFSGSAYIVVSRVYVGVLOQNEIOLKINEMAQDTTSP : 318
 IGnT : NITPGVLEPPAIGETKYVQEHLE--KELSYVIRTTALKEPEENIMETGNAYIVASRDEVQHLKNPSOOLIEVVKDTTSP : 294

*

C2GnT3 : DEHWATIIRVPGIPGEISRS-AQDVSDLOSCTRUVKNNYFEG-----FFYBSCTGSHLSRVCILYGAEIRWHIKDGEWEANKFP : 413
 C2GnT2 : DEHLWATILRARMPCGSVPHHPDIDSDNTSIRREVKTGHEGIDKGAHYAPCSGIHQRAICVMGAGDINMTQONHHLANKFP : 413
 C2GnT1 : DEYLWATILRPEVPGSILPASHDLSDMQAVARFVKTMFESVSKGAPYEPDGVHVRVSVCLIFGAGDINMTQONHHLANKFP : 403
 IGnT : DEHWMVTLNRPVPCGSMPEA--SWT-----GNIRAIKTSMDMED-----RHGGCHGHYVHGCILYGNGLKWAIVNSPSEANKKE : 367

*

C2GnT3 : SKVDIILKCLAEKLEEQORDWITTPSEKTFMDRNLTTTS : 453
 C2GnT2 : PVDDNATLCLAEYLYA-----AIYGT----- : 438
 C2GnT1 : VDVLFLCLDEHLHF-----AETLKH----- : 428
 IGnT : LNTYLTVECLLRHR-----TLNQSTAIQPSWYF-- : 400

Fig. 2

www.nature.com/scientificdata/

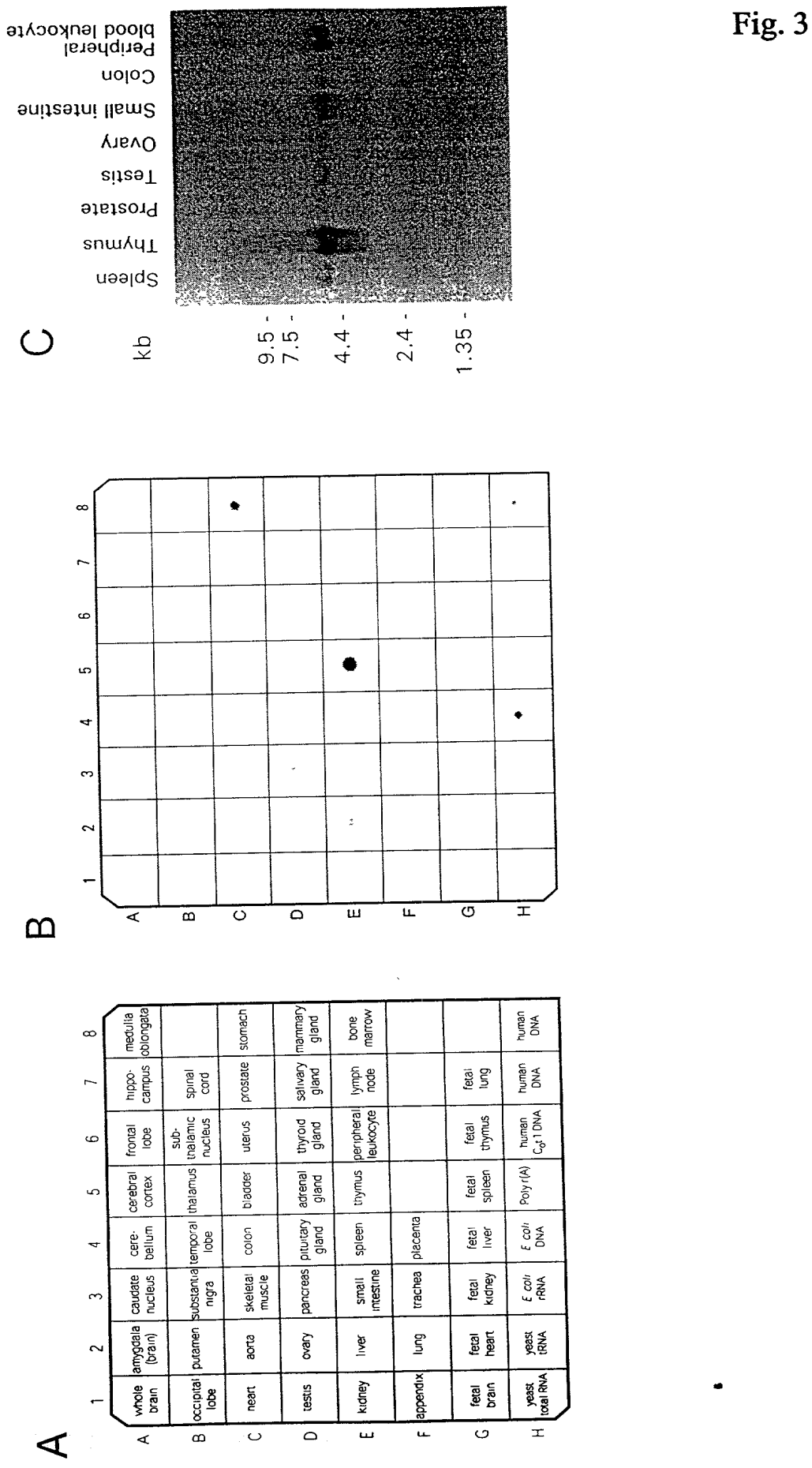
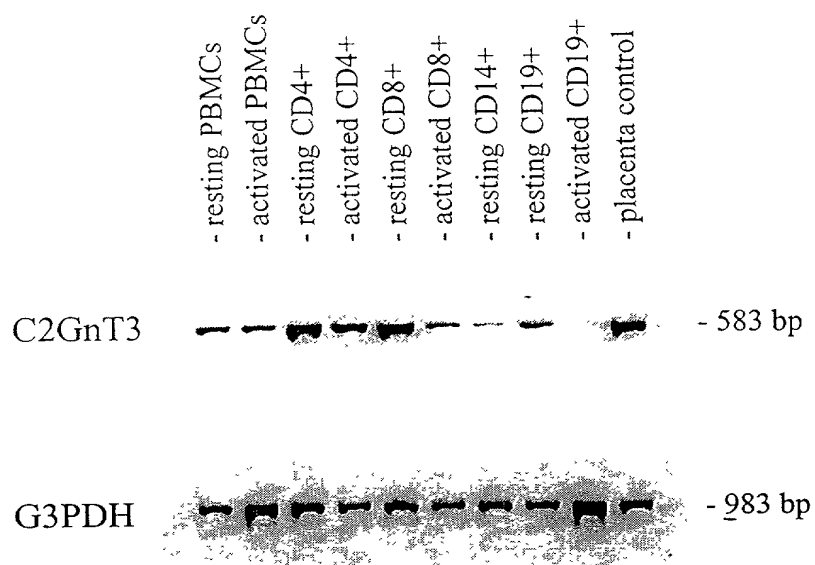
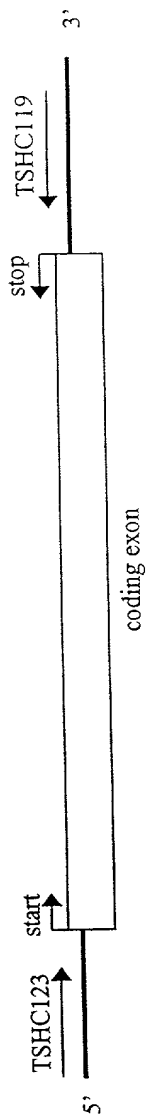


Fig. 3

Fig. 4



C2GnT3



TSHC123: 5'-GGGCAGCATTTGCCCTAGTATG-3'

TSHC119: 5'-GATCTCTGATTGGCTCAGTG-3'

Fig. 5